ICASE

ON THE CONVERGENCE PROPERTIES OF THE DENSITY

MATRIX DIRECTED GENERAL SECOND-ORDER MCSCF ALGORITHM

Byron H. Lengsfield, III

and

Charles W. Bauschlicher, Jr.

Report No. 80-19

July 14, 1980

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

ON THE CONVERGENCE PROPERTIES OF THE DENSITY MATRIX DIRECTED GENERAL SECOND-ORDER MCSCF ALGORITHM

Byron H. Lengsfield, III[*,†]
NASA Langley Research Center

and

Charles W. Bauschlicher, Jr.[†]
Institute for Computer Applications in Science and Engineering

ABSTRACT

The density matrix directed second order MCSCF algorithm is reviewed. The rate of convergence is discussed and several examples are given, showing the importance of including as many CI roots as possible.

---

[*] A contribution by B. H. Lengsfield, III. NRC-NASA Research Associate, 1978-1980. Present address: IBM, 5600 Cottle Road, San Jose, CA 95193.

[†] Work was supported under NASA Contracts No. NAS1-15810 and NAS1-14472 while in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

ON THE CONVERGENCE PROPERTIES OF THE DENSITY MATRIX DIRECTED GENERAL SECOND-ORDER MCSCF ALGORITHM

Byron H. Lengsfield, III*,†
Laser and Spectroscopy Branch
NASA Langley Research Center, Hampton, VA 23665

Charles W. Bauschlicher, Jr.†
Institute for Computer Applications in Science and Engineering
NASA Langley Research Center, Hampton, VA 23665

The density matrix directed (DMD) second order MCSCF algorithm[1] (see also ref. 2-6) is based on the fact that one can construct the Hessian and the gradient of the energy expression from the unique elements of one and two particle density matrices. With the recent development of the Unitary Group CI method,[7,8,9] density matrix elements can be obtained even for very large MCSCF problems without excessive computational effort.[10,11] Alternatively, one can obtain these density matrix elements by sorting a conventional CI formula tape,[1] and this technique has proven to be quite practical for traditional MCSCF problems.

The energy of a general MCSCF wavefunction

$$\Psi^I = \sum_k^N C_k^I \phi_k \tag{1}$$

can be expressed as follows

$$E^I = \sum_{i\geq j} D_{ij}^I (\phi_i |h| \phi_j) + \sum_{\substack{i\geq j \\ (ij)\geq(k\ell)}} \sum_{k\geq \ell} D_{ijk\ell}^I (\phi_i \phi_j | g | \phi_k \phi_\ell) \tag{2}$$

where

$$D_{ij}^I = \sum_{o\geq p}^N C_o^I C_p^I SF_{op}^{ij} \tag{3a}$$

$$D_{ijk\ell}^I = \sum_{o\geq p}^N C_o^I C_p^I SF_{op}^{ijk\ell} \tag{3b}$$

and $SF_{k\ell}^{ij}$ are structure factors. Variations in eqn. (2) are introduced by means of exponential unitary transformations[12] of the molecular orbitals (ϕ_i)

$$\tilde{U}_{mo} = e^{\tilde{\Delta}} = \tilde{1} + \tilde{\Delta} + \tfrac{1}{2}\tilde{\Delta}^2$$

$$\tilde{\Delta} = \begin{pmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ -\Delta_{13} & -\Delta_{23} & 0 & \Delta_{34} \\ -\Delta_{14} & -\Delta_{24} & -\Delta_{34} & 0 \end{pmatrix}$$

$$\tilde{\Delta} = \begin{pmatrix} \tilde{P}_1 \\ \tilde{P}_2 \\ \tilde{P}_3 \\ \tilde{P}_4 \end{pmatrix} \qquad \begin{matrix} \text{Full} = 2 \\ \text{Part} = 2 \end{matrix}$$

$$\Psi = c_1(f_1^2 f_2^2 p_1^2) + c_2(f_1^2 f_2^2 (p_1 p_2))$$

$$\tilde{\Delta}^2 = \begin{pmatrix} (-\Delta_{13}^2 - \Delta_{14}^2) & (-\Delta_{13}\Delta_{23} - \Delta_{14}\Delta_{24}) & (-\Delta_{14}\Delta_{34}) & (\Delta_{13}\Delta_{34}) \\ (-\Delta_{13}\Delta_{23} - \Delta_{14}\Delta_{24}) & (-\Delta_{23}^2 - \Delta_{24}^2) & (-\Delta_{24}\Delta_{34}) & (\Delta_{23}\Delta_{34}) \\ \vdots & \cdot & \cdot & \cdot \end{pmatrix}$$

$$\tilde{\Delta}^2 = \begin{pmatrix} \tilde{Q}_1 \\ \tilde{Q}_2 \\ \tilde{Q}_3 \\ \tilde{Q}_4 \end{pmatrix} \tag{4}$$

and the CI vectors (C^I).

$$\tilde{U}_{CI} = e^{\tilde{\gamma}} = \tilde{1} + \tilde{\gamma} + \tfrac{1}{2}\tilde{\gamma}^2$$

$$\tilde{\gamma} = \begin{pmatrix} 0 & \gamma_{12} & 0 & 0 \\ -\gamma_{12} & 0 & \gamma_{23} & \gamma_{24} \\ 0 & -\gamma_{23} & 0 & 0 \\ 0 & -\gamma_{24} & 0 & 0 \end{pmatrix} \quad \text{— } P_2$$

$$\tilde{\gamma}^2 = \begin{pmatrix} -\gamma_{12}^2 & 0 & (\gamma_{12}\gamma_{23}) & (\gamma_{12}\gamma_{24}) \\ 0 & (-\gamma_{12}^2-\gamma_{23}^2-\gamma_{24}^2) & 0 & 0 \\ \gamma_{12}\gamma_{23} & 0 & -\gamma_{23}^2 & -\gamma_{23}\gamma_{24} \\ \gamma_{12}\gamma_{24} & 0 & -\gamma_{23}\gamma_{24} & -\gamma_{24}^2 \end{pmatrix} \quad \text{— } Q_2 \tag{5}$$

where $\tilde{P}_1$ and $\tilde{Q}_1$ are row generators which define the non-redundant orbital and CI mixings at first and second order respectively, (i.e., $\tilde{P}_1$

---

*a contribution by B. H. Lengsfield, III. NRC-NASA Research Associate, 1978-1980. Present address: IBM, 5600 Cottle Road, San Jose, CA 95193.

†Work was supported under NASA Contracts No. NAS1-15810 and NAS1-14472 while in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

generates the non-zero elements of the i^{th} row of $\tilde{\Delta}$ and $\tilde{Q}_i$ generates the non-zero elements of the i^{th} row of Δ^2). The energy expression is then expanded to second order in terms of the non-redundant variables of the generators of the two unitary transformations.

The contributions to gradient and Hessian can be obtained quite simply in terms of the $\tilde{P}$ and $\tilde{Q}$ row generators.

GRADIENT TERMS

ORBITAL MIXINGS

$$\tilde{P}_1 \overset{I}{D}_{ijk\ell} (\phi i\, \phi j \,|g|\, \phi k\, \phi \ell) = \overset{I}{D}_{ijk\ell} \sum_z{}' (\phi z\, \phi j \,|g|\, \phi k\, \phi \ell)\ \Delta_{iz}\ ph(i, z)$$

$$ph(i, z) = 1 \text{ if } z > i$$
$$= -1 \text{ if } z < i \qquad (6)$$

C.I. MIXINGS

$$\tilde{P}_I \overset{I}{D}_{ijk\ell} (\phi i\, \phi j \,|g|\, \phi k\, \phi \ell) = (\phi i\, \phi j \,|g|\, \phi k\, \phi \ell) \sum_J{}' \gamma_{IJ}\ ph(I, J)$$

$$+ \sum_{o\geq p}^{N} (C_o^I C_p^J + C_o^J C_p^I)\ SF_{op}^{ijk\ell} \qquad (7)$$

HESSIAN TERMS

ORBITAL-ORBITAL INTERACTIONS

(1) $\tilde{P}i\ \tilde{P}j\ D_{ijk\ell} (\phi i\, \phi j \,|g|\, \phi k\, \phi \ell)$

$$= D_{ijk\ell} \sum_Y{}' \sum_Z{}' (\phi z\, \phi y \,|g|\, \phi k\, \phi \ell)\ \Delta_{iz}\ \Delta_{jy}\ ph(i,z)\ ph(j,y)$$

(2) $\tilde{Q}i\ D_{ijk\ell} (\phi i\, \phi j \,|g|\, \phi k\, \phi \ell)$

$$= D_{ijk\ell} \sum_q^{N} (\phi q\, \phi j \,|g|\, \phi k\, \phi \ell) \sum_m^{M} \Delta_{im}\ \Delta_{qm}\ ph(i,m)\ ph(m,q)$$

$$\Delta_{im}\ \Delta_{jm} \rightarrow -1$$

$$\Delta_{im}\ \Delta_{mq} \rightarrow +1 \qquad (8)$$

C. I.-ORBITAL INTERACTIONS

$$\tilde{P}_I\ \tilde{P}i\ \overset{I}{D}_{ijk\ell} (\phi i\phi j \,|g|\, \phi k\phi \ell)$$

$$= \left\{\tilde{P}_I\ \overset{I}{D}_{ijk\ell}\right\} \left\{\tilde{P}i\ (\phi i\phi j \,|g|\, \phi k\phi \ell)\right\} \qquad (9)$$

$$\therefore \vec{P}_{CI} = \vec{P}_{MO}^{\dagger} = \left(\partial^2 E' / \partial \gamma \partial \Delta\right)$$

C. I. - C. I. INTERACTIONS

$$\tilde{P}_I \tilde{P}_I\ D_{ijk\ell} (\phi i\phi j \,|g|\, \phi k\phi \ell)$$

$$= (\phi i\phi j \,|g|\, \phi k\phi \ell) \sum_{o\geq p} SF_{o,p}^{ijk\ell} \sum_J^{N} \sum_K^{N} C_o^J C_p^K\ \gamma_{IJ}\gamma_{IK}\ ph(I,J)\ ph(I,K)$$

$$= \sum_J^{N} \sum_K^{N} \langle C^J \,|\, \tilde{H} \,|\, C^K \rangle\ \gamma_{IJ}\gamma_{IK}\ ph(I,J)\ ph(I,K) \qquad (10)$$

RECALL THERE IS ONLY ONE Q TYPE C. I. TERM

DIAGONAL ELEMENT OF THE C. I. - C. I. HESSIAN

$$\left\{\langle C^J \,|\, \tilde{H} \,|\, C^J \rangle - \langle C^I \,|\, \tilde{H} \,|\, C^I \rangle\right\} \gamma_{IJ}^2$$

$$\longrightarrow \{E_J - E_I\} \qquad \text{if} \quad \tilde{H}\ C^k = E_K\ C^k \qquad (11)$$

The DMD-MCSCF algorithm is outlined in Figure 1.

MCSCF ALGORITHM

1. CONFIGURATION GENERATION
2. C. I. FORMULA GENERATION
3. D. M. SORT OF THE C. I. FORMULAS
4. HESSIAN FORMULA GENERATION
5. MCSCF ITERATIONS
   a. ORTHOGONALIZATION OF THE ORBITALS
   b. INTEGRAL TRANSFORMATION
   c. CONSTRUCTION AND DIAGONALIZATION OF THE HAMILTONIAN
   d. CONSTRUCTION OF THE HESSIAN AND THE GRADIENT, FOLLOWED BY THE SOLUTION OF THE LINEAR EQUATIONS
   e. APPROXIMATE UNITARY TRANSFORMATION OF THE ORBITALS

Figure 1.

As noted by Siegbahn, Heilberg, Roos, & Levy[5] (see ref. 2), the density matrix elements which involve only core orbitals need not be treated separately and various two electron density matrix elements in which two of the indices refer to core orbitals may also be grouped together. The Fock operator expressions employed by Siegbahn, et al. in the construction of the gradient may also be used to advantage in the construction of the CI-orbital portion of the Hessian. Furthermore, the density matrix elements whose indices refer only to core orbitals do not contribute to the CI-orbital portion of the Hessian.

Finally we note several advantages obtained by employing the eigenvectors of the Hamiltonian in this technique. First one need not include all of the CI vectors in the expansion of the energy to obtain a variational algorithm. This, of course, allows one to address much larger MCSCF problems. Second, the CI-CI portion of the Hessian and the CI terms in the gradient are diagonal and zero respectively. This allows the equations for the elements of the unitary generators to be simplified. Finally, redundant variables are not always easily identified in large MCSCF calculations. However, when the root of the Hamiltonian which is being optimized is an eigenvector of the Hamiltonian, redundant variables generally give rise to zeros in the gradient and can be easily detected.

There are perhaps two points of primary interest in evaluating the convergence properties of this alogrithm. The first point being the behavior of this procedure when poor starting orbitals are employed. The second point is concerned with the number of CI vectors which must be included in large problems to obtain a reasonable rate of convergence.

In a large number of MCSCF problems one does not possess a very good choice of starting orbitals. This is especially true if the MCSCF wavefunction contains several configurations which differ from one another by (spin orbital) single excitations. In this case the Hessian may posses very small and even negative eigenvalues. Various means have been proposed to shift the eigenvalues (or alternatively the diagonal elements)[13] of the Hessian or to take a step in the direction indicated by the eigenvector which corresponds to the negative eigenvalue of the Hessian.[2] These techniques often yield disappointing results especially when applied to problems where the MCSCF reference contains CSF's which differ by a single excitation.[1] Instead, a super CI technique has been found to possess a much larger radius of convergence[1] as demonstrated in Table I. In this method one constructs an augmented Hessian matrix [1,14,15] in analogy with the Singles Hamiltonian

SUPER - C. I. APPROACH

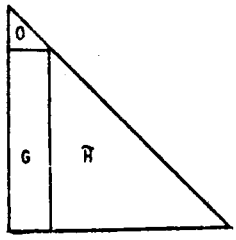

Figure 2.

constructed in Generalized Brillouin Theorem algorithms. It is important to note that a density matrix oriented approach allows this matrix to be constructed in a much more efficient manner than the traditional contraction type procedures.[16-20] This method is particularly attractive as quadratic convergence[1] is very often obtained (when all CI coupling terms are included) in this procedure when a reasonable set of orbitals has been obtained. Moreover, a simple scheme can be devised to correct the eigenvector of this matrix, when it is not dominated by the MCSCF reference and further increase the radius of convergence attained by this technique.

In the following tables the results of several MCSCF calculations are presented in which the number CI vectors included in the Hessian has been varied.

TABLE I. $BeO^{a,b}$ 3 CSF SUPER CI ALGORITHM

Iteration	Energy (a.u.)	ΔE	Δ^c
1[d]	-89.424317	-	2.56 E-1
2	-89.466285	4.20 E-2	2.29 E-1
3	-89.495331	2.90 E-2	1.13 E-1
4	-89.503767	8.44 E-3	6.62 E-2
5	-89.505765	2.00 E-3	1.53 E-2
6	-89.506026	2.60 E-4	5.07 E-4
7	-89.506033	7.99 E-6	6.57 E-7
8	-89.506034	1.13 E-8	1.42 E-12

a) core $4\sigma^2\pi^4$, core $4\sigma5\sigma\pi^4$ and core $4\sigma^2\pi^3\pi$

b) Bauschlicher-Yarkony Basis, J. Chem. Phys. 72, 1138 (1980)

c) $\Delta \equiv \Sigma \Delta_{ij}^2$, see eqn. (4)

d) Damping performed, the Hessian possessed two negative eigenvalues. SCF starting guess employed.

TABLE II. APPROXIMATE SUPER CI[a]

Iteration	Energy (a.u)	ΔE	Δ
1	-89.424317	-	.162
2	-89.462099	3.78 E-2	.137
3	-89.487145	2.50 E-2	8.36 E-2
4	-89.497525	1.04 E-2	2.93 E-2
5	-89.501648	4.12 E-3	1.40 E-2
6	-89.503637	1.99 E-3	8.40 E-3
7	-89.504713	1.08 E-3	5.00 E-3
8	-89.505311	5.98 E-4	2.82 E-3
9	-89.505642	3.31 E-4	1.54 E-3
10	-89.505824	1.82 E-4	8.19 E-4

a) The second order CI terms were not included in this calculation.

The importance of including or excluding a particular CI vector can be placed on a more quantitative basis by means of the simple perturbation arguments presented below.

Consider the Newton-Raphson linear equations induced by a two CSF, two orbital problem,

$$\begin{pmatrix} B & C \\ C & A \end{pmatrix}\begin{pmatrix} \Delta \\ \gamma \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix} \qquad (12)$$

where B, C and A represent the orb.-orb., CI-orb., and CI-CI portions of the Hessian respectively. Δ and γ are the unique elements of the generators of the unitary transformations and g is the orbital gradient (the CI gradient is zero by virture of the fact that the secular is assumed to have been solved on the preceding iteration). We then have,[2]

$$\left(B - \frac{C^2}{A}\right)\Delta = g \qquad (13)$$

$$\Delta = \left(\frac{g}{B - \frac{C^2}{A}}\right)$$

$$= \frac{g}{B}\left(\frac{1}{1 - \frac{C^2}{AB}}\right) \qquad (14)$$

$$\simeq \frac{g}{B}\left(1 + \frac{C^2}{AB} + \cdots\right) .$$

(Recall $A \alpha E^2 - E^1$ where 1 is the root being optimized.) This perturbation may be generalized to account for the interaction between a particular CI vector and all of the orbital mixings. For the purposes of this study it suffices to consider these interactions a sum of decoupled two by two problems and monitor the largest perturbation term (C^2/AB) associated with the last CI vector to be included in the problem. These terms are also presented in Tables III and IV. It is interesting to note that this perturbation term differs from the second-order perturbation expression obtained by Das[21] in that the energy difference appearing in the denominator is weighted a diagonal element from the orbital section of the Hessian.

The results of a number of MCSCF calculations are summarized in the following tables.

TABLE IIIa. 3 CSF BeO CALCULATION WITH ONE CI ROOT EXCLUDED

	E	ΔE	Δ
1	-89.424317	-	2.61 E-1[a]
2	-89.465898	4.16 E-3	2.28 E-1
3	-89.494027	2.81 E-2	1.03 E-1
4	-89.502397	8.37 E-3	2.82 E-2
5	-89.504267	1.87 E-3	8.63 E-3
6	-89.505109	8.41 E-4	3.71 E-3
7	-89.505550	4.41 E-4	1.86 E-3
8	-89.505782	2.32 E-4	9.58 E-4
9	-89.505903	1.21 E-4	4.92 E-4
10	-89.505967	6.3 E-5	2.51 E-4

a) Damping employed this iteration

TABLE IIIb. BeO 3 CSF MCSCF

Root	Largest Perturbation Term	Sum of Perturbation Contributions
1	.18	.72
2	.31	1.48

TABLE IV. FULL VALENCE $^1\Sigma^+$ STATE OF MgO[a,b]

Number of CI Roots	0		30		90	
Iteration	ΔE	Δ	ΔE	Δ	ΔE	Δ
1	-	9.E-5	-	1.E-4	-	2.E-4
2	-3.E-6	6.E-6	-4.E-6	1.E-5	-5.E-6	3.E-6
3	-9.E-7	2.E-6	-5.E-7	2.E-6	-8.E-7	1.E-7
Final Energy	-274.514267		-274.514268		-274.514268	

a) Yoshimine-McLean molecular DZP Slater basis with a diffuse 3s-function (.855) on Mg. 142 CSF's in C_{2v}

b) The maximum perturbation term (.43) was obtained from the 22 eigenvector of the Hamiltonian.

The perturbation contributions of the higher eigenvectors of the Hamiltonian are often larger than the perturbation estimate of the contribution of many of the lower roots of the Hamiltonian. While the inclusion of a few of the CI vectors in the varitional problem can dramatically affect the convergence characteristic of this MCSCF algorithm far from convergence. One can not expect any substantial advantage from this procedure near convergence for a general (containing single excitations) MCSCF wavefunction.

The DMD-MCSCF algorithm provides a simple and efficient means for constructing the Hessian and the gradient of a general MCSCF energy expression. The studies reported in the paper and recent work on MgO and BeO[22] indicates that this method is capable of providing the convergence behavior needed to perform practical quantum chemical calculations.

References

[1] B. H. Lengsfield, III, J. Chem. Phys. xx, xxx (1980).

[2] L. G. Yaffe and W. A. Goddard, III, Phys. Rev. A 13, 1682 (1976).

[3] E. Daalgard and P. Jorgensen, J. Chem. Phys. 69, 3833 (1979).

[4] C. C. J. Roothaan, J. Detrich and D. G. Hopper, Int. J. Quantum. Chem. Symp. 13, 93 (1979).

[5] P. Siegbahn, A. Heilberg, B. Roos and B. Levy, Physica Scripta 21, 323 (1980).

[6] P. Jorgensen and D. Yeager, J. Chem. Phys. 71, 757 (1979).

[7] J. Paldus and M. J. Boyle, Physica Scripta, 21 295 (1980), and references therein.

[8] I. Shavitt, Int. J. Quant. Chem. Symp. 11, 131 (1977), 12, 5 (1978).

[9] B. R. Brooks and H. F. Schaefer, III, J. Chem. Phys. 70, 5092 (1979)

[10] B. R. Brooks, W. D. Laidig, P. Saxe, N. C. Handy, and H. F. Schaefer, III, Physica Scripta 21, 312 (1980).

[11] B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi and H. F. Schaefer, III, J. Chem. Phys. 72, 4652 (1980).

[12] K. M. Sando and J. E. Harriman, J. Chem. Phys. 47, 180 (1967).

[13] V. A. Kuprevich and V. E. Klymeno, Int. J. Quantum Chem. 12, 327 (1977).

[14] T. C. Chang and W. H. E. Schwartz, Theor. Chem. Acta 44, 45 (1977).

[15] F. W. Bobrowicz and W. A. Goddard, III, pp. 79-127, Vol. 3, Modern Theoretical Chemistry, ed. H. F. Schaefer, III (Plenum, New York, 1979).

[16] B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968).

[17] F. Grein and A. Banarjee, Int. J. Quantum Chem. Symp. 9, 147 (1975).

[18] F. Grein and T. C. Chang, Chem. Phys. Lett. 12, 44 (1971).

[19] P. J. A. Ruttink and J. H. van Lenthe, Theor. Chem. Acta 44, 97 (1977).

[20] L. M. Cheung, T. S. Elbert, and K. Ruedenberg, Int. J. Quantum Chem. 16, 1069 (1979).

[21] G. Das, J. Chem. Phys. 58, 5104 (1973).

[22] C. W. Bauschlicher, Jr., B. H. Lengsfield, III, D. M. Silver, and D. R. Yarkony, to be published.